Resolving the universe with multifractals

Cristiano G. Sabiu(1), Luìs F. A. Teodoro(2), Martin A. Hendry(2)

(1) *Institute of Cosmology & Gravitation, University of Portsmouth, England*
(2) *Astronomy & Astrophysics Group, University of Glasgow, Scotland*

Summary. — We present a new method for dealing with geometrical selection effects in galaxy surveys while using a multifractal framework. The power of multifractal analysis lies in its connection to higher order moments, in that it not only probes clustering on different scales but also different densities. Therefore any incompleteness issues must be correctly addressed before blindly applying this technique to real survey data.

PACS `98.80.Es` – Observational cosmology.
PACS `02.70.Rr` – General statistical methods.

1. – Introduction

There has been much work done in trying to characterise the galaxy distribution using multifractals [5, 6, 8]. One of the main problems which arises in a multifractal analysis is how to deal with observational selection effects, *i.e.* 'masks' in the survey region and a geometric boundary to the survey itself. In this paper we will introduce a *volume* boundary correction which is rather similar to the approach developed by [9].

In §**2** the methodology is constructed to analyse data within a multifractal framework, this then leads us to several problems in galactic surveys which must be accounted for. We will then discuss previous attempts at correcting for edge effects in §**3** and show that there is still room for improvement. Then a comparison is drawn between these different methods using a toy model in §**4**.

2. – Multifractal Formalism

In this analysis we will adopt the procedure layed out in [4] to determine the Rényi (Generalised) dimensions of a point set embedded in a three-dimensional Euclidean space.

---

(1) email: cris.sabiu@port.ac.uk

The probability of a galaxy, j, being within a sphere of radius r centred on galaxy i is,

$$(1) \qquad p_i(r) = \frac{n_i(r)}{N} = \frac{1}{N} \sum_{j=1}^{N} \Phi(|r_i - r_j| - r).$$

Here $n_i(r)$ is the number of galaxies within radius r, N is the total number of galaxies and $\Phi(x)$ is the Heaviside step function. Equation (1) can then be related to the partition sum via the correlation algorithm introduced in [2]

$$(2) \qquad Z(q,r) = \frac{1}{M} \sum_{i=1}^{M} [p_i(r)]^{q-1} \propto r^{\tau(q)}.$$

In this case M is the number of counting spheres and q defines the generalised dimension we are investigating. $\tau(q)$ is the scaling exponent, which is then related to the infinite set of dimensions through,

$$(3) \qquad D_q = \frac{\tau(q)}{q-1}$$

3. – Boundary Corrections

Firstly, and most obviously, there is the *deflation* method. This simply makes no correction at all, it only allows counting spheres which lie completely within the survey. This dramatically reduces the distance out to which our estimators can probe and also reduces the number of large spheres we can average over, leading to a sort of cosmic variance.

The *deflation* method therefore wastes a great amount of data especially when used with a 'pencil beam' or 'thin fan' survey. It would be more useful if our estimators could sample all of the survey without worrying about its geometry. This however leads to counting cells exceeding the survey boundaries and consequently missing galaxies [3].

To deal with partially filled/empty cells we could use the *capacity* correction. This method populates the missing regions with galaxies (R2 in Fig.1). At first this may seem like an appropriate coarse of action but we are then forced to decide on the right distribution of galaxies to fill the void. Previous attempts at analysing redshift surveys have used number densities like, $n \propto r^3$ (*e.g.* [1]). This is of coarse assuming an answer to the question we pose, as the exponent is fixed to be 3. It is then not surprising that a transition to homogeneity has been reported by groups using this approach.

In Fig.1 the counting sphere has exceeded the geometrical boundary of the survey. The number of galaxies counted in the sphere of radius r is depleted which leads to $p_i(r)$ being reduced through Eq. (1). To solve this problem we could either add galaxies to the missing region like the *capacity* correction or we could somehow modify the volume (hence the name!). We can cast expression (1) as,

$$(4) \qquad p_i(r) = \frac{V_i(r)\rho^\star(r)}{N} = \frac{V_i(r)}{V_i^\star(r)} \cdot \frac{n_i^\star(r)}{N},$$

where now we have V being the real volume of the sphere and $V^\star$ is the reduced volume. We have also introduced the reduced density as an intermediary step which need not be

calculated and related this with a reduced volume and number count, $n_i^\star$. On its own this method can be visualised in Fig.1, as giving the missing region 2 the same density as region 1. This would be wrong if density varies with distance, so that $\rho_{R1} \neq \rho_{R2}$. To overcome this problem we assume *only* that the density does not vary with θ or ϕ i.e. the universe is isotropic and hence Eq. (4) will hold for fixed r. So to apply this method to a galaxy survey we must count in spherical shells, correcting as we go and then integrate up the shells at the end. This method is shown in Fig.2. The shells are individually corrected and summed according to,

$$(5) \qquad p_i(r) = \sum_{r=0}^{r} \alpha_i(r) \frac{n_i^\star}{N}$$

Where $\alpha_i(r) \equiv V/V^\star$, this is the enhancement factor of the ith shell at radius r and has value ≥ 1. The main advantage of this method is that makes full use of the data. The deflation method on the other hand throws away a lot of useful data and hence increase the errors in their results.

4. – TESTING THE CORRECTIONS

In this section we will begin by applying the different corrections to a toy fractal model. This model can be used to compare and contrast the differing methods since it has an analytically determined dimension. We also set up the fractal distributions to mimic real galaxy survey by being of comparable size (volume and number count). Then the *volume* correction will be used to analyse in detail the distribution of particles produced from N-body simulations.

4˙1. *MULTIPLICATIVE CASCADE*. – To test further the power of our multifractal analysis we construct a hierarchical clustering model in two dimensions. This model is called the Multiplicative Cascade and was first discussed in Meakin (1987) [7], it has the advantage over the Levy Flight in that the whole of it Dq curve can be calculated analytically.

The fractal is constructed as follows: The space is split into four equal parts, each part is then assigned a probability from the set $\{p_1, p_2, p_3, p_4\}$ without replacement. Where $p_i \in [0,1]$. Each subspace is then divided again and assigned probabilities randomly from the same set and this is continued to the Nth level.

At the Nth level the probability of a cell being occupied is the product of the cell's p_i and its parents and ancestors up to level 1 *i.e.* all the cells above it. In constructing this model down to level 8 we produce a 4^8 array of cells each with its own probability. To then place particle in the space we invoke a Monte Carlo rejection scheme. Choosing x and y coordinates randomly we simply test if a random number between 0 and 1 is less or greater than the cell probability. We typically dust the probability density field with 5,000 points in a space of 256 x 256 h^{-1} Mpc. The density field can be visualized in Fig.3.

It can be shown [5] that as $L \to \infty$,

$$(6) \qquad D_q = \frac{\log_2 (f_1^q + f_2^q + f_3^q + f_4^q)}{1 - q},$$

Model	p_1	p_2	p_3	p_4
I	1	1	1	0
II	1	0.75	0.75	0.5
III	1	0.5	0.5	0.25

TABLE I. – *Parameters of the Multiplicative Cascade models*

where

$$(7) \qquad\qquad f_i = \frac{p_i}{\sum_i p_i}.$$

Fig.4 shows three fractal models with their measured and theoretical Renyi Dimensions.

5. – **Conclusions & Future Work**

Our 'volume' correction and implementation of the Multifractal analysis successfully recovers the underlying Generalised dimensions of the multifractal distributions. It does a better job at correcting for the survey boundary than the methods discussed earlier and has the potential to work with much more complicated incomplete geometries. We hope soon to apply this method of analysis to real redshift surveys with the hope of uncovering the true hierarchical nature of the galaxy distribution. For a more detailed discusison see Sabiu 2007 [10].

<center>* * *</center>

REFERENCES

[1] S. Borgani, V. J. Martinez, M. A. Perez, and R. Valdarnini. Is there any scaling in the cluster distribution? *ApJ*, 435, November 1994.
[2] P. Grassberger and I. Procaccia. Estimation of the Kolmogorov entropy from a chaotic signal. *Phys. Rev. A*, 28, October 1983.
[3] S. Hatton. Approaching a homogeneous galaxy distribution: results from the Stromlo-APM redshift survey. *MNRAS*, 310:1128–1136, December 1999.
[4] H. G. E. Hentschel and I. Procaccia. Fractal nature of turbulence as manifested in turbulent diffusion. *Phys. Rev. A*, February 1983.
[5] B. J. T. Jones, V. J. Martinez, E. Saar, and J. Einasto. Multifractal description of the large-scale structure of the universe. *ApJ Lett.*, 332, September 1988.
[6] V. J. Martinez, B. J. T. Jones, R. Dominguez-Tenreiro, and R. van de Weygaert. Clustering paradigms and multifractal measures. *ApJ*, 357:50–61, July 1990.
[7] P. Meakin. Diffusion-limited aggregation on multifractal lattices: A model for fluid-fluid displacement in porous media. *Phys. Rev. A*, 36:2833–2837, September 1987.
[8] J. Pan and P. Coles. Large-scale cosmic homogeneity from a multifractal analysis of the PSCz catalogue. *MNRAS*, 318:L51–L54, November 2000.

[9] J. Pan and P. Coles. Boundary corrections in fractal analysis of galaxy surveys. *MNRAS*, March 2002.

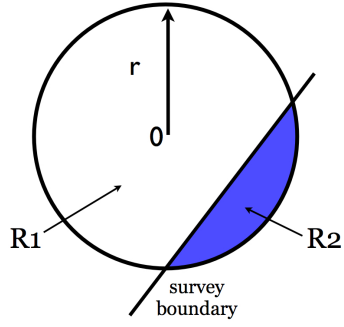[10] Cristiano Sabiu. Probing the large-scale homogeneity of the universe with galaxy redshift surveys, 2007.

Fig. 1. – A counting sphere centred on galaxy O with radius r. Regions R1 & R2 are inside and outside the survey respectively.
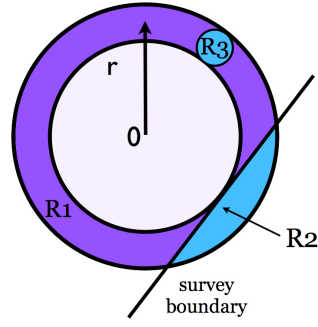

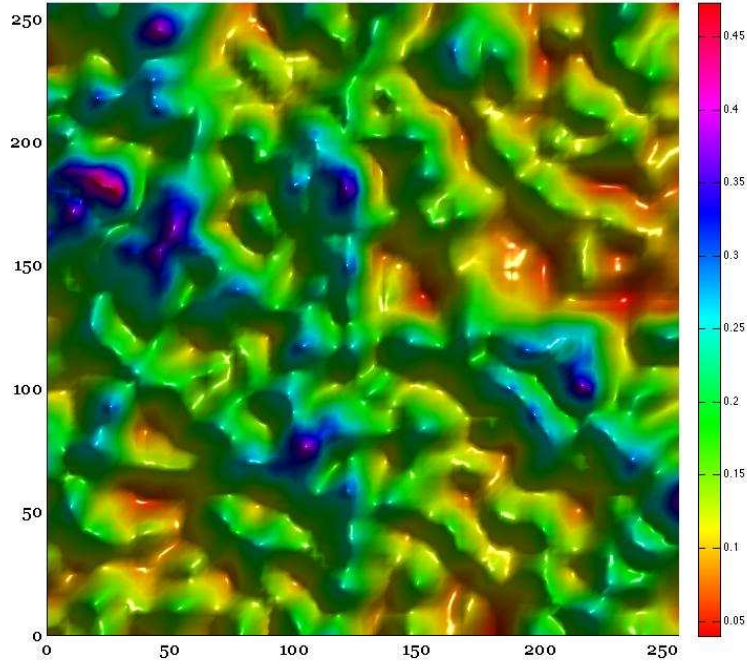Fig. 2. – Region R1 is a shell inside the survey, R3 is a unobserved hole & R2 is the boundary of the survey.

Fig. 3. – This is an example of our probability density field as explained in §**4**˙1.
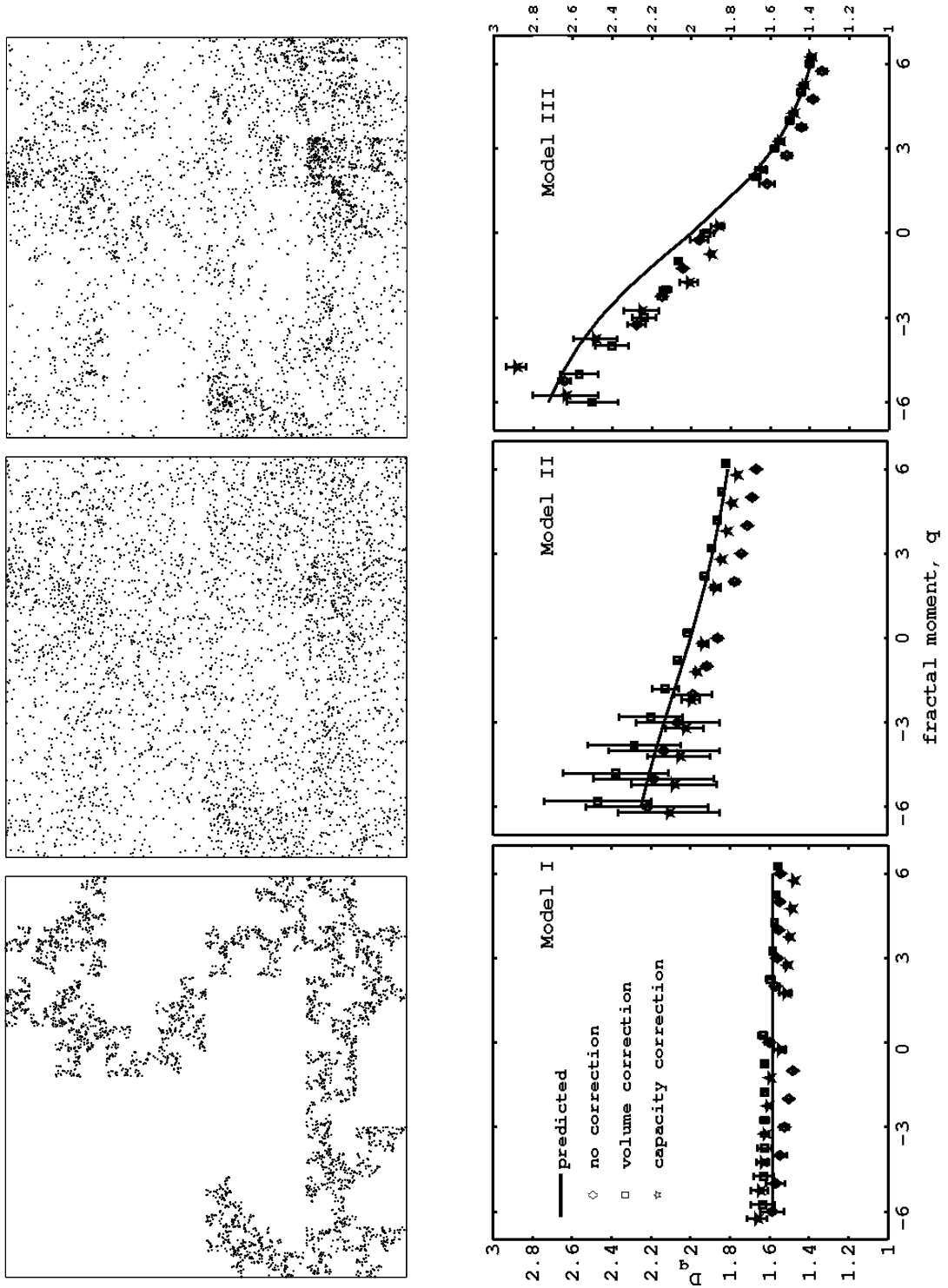
Fig. 4. – Three realizations of our model fractal. Model I is a monofractal with uniform dimension $\log(3)/\log(2)$. Models II and III are multifractal distributions and are statistically inhomogeneous. The solid line is the theoretical model for the D_q curve.